



Estyn

Rhagoriaeth i bawb - Excellence for all

Arolygiaeth Ei Mawrhydi dros Addysg
a Hyfforddiant yng Nghymru

Her Majesty's Inspectorate
for Education and Training in Wales

Evaluation of the arrangements to assure the consistency of teacher assessment in the core subjects at key stage 2 and key stage 3



BUDDSODDWR MEWN POBL
INVESTOR IN PEOPLE



MARCH 2010

The purpose of Estyn is to inspect quality and standards in education and training in Wales. Estyn is responsible for inspecting:

- ▲ nursery schools and settings that are maintained by, or receive funding from, local authorities (LAs);
- ▲ primary schools;
- ▲ secondary schools;
- ▲ special schools;
- ▲ pupil referral units;
- ▲ independent schools;
- ▲ further education;
- ▲ adult community-based learning;
- ▲ youth support services;
- ▲ youth and community work training;
- ▲ LAESs;
- ▲ teacher education and training;
- ▲ work-based learning;
- ▲ careers companies;
- ▲ offender learning; and
- ▲ the education, guidance and training elements of the Department for Work and Pensions funded training programmes.

Estyn also:

- ▲ provides advice on quality and standards in education and training in Wales to the National Assembly for Wales and others; and
- ▲ makes public good practice based on inspection evidence.

Every possible care has been taken to ensure that the information in this document is accurate at the time of going to press. Any enquiries or comments regarding this document/publication should be addressed to:

Publication Section

Estyn

Anchor Court

Keen Road

Cardiff

CF24 5JW or by email to publications@estyn.gov.uk

This and other Estyn publications are available on our website: www.estyn.gov.uk

© Crown Copyright 2010: This report may be re-used free of charge in any format or medium provided that it is re-used accurately and not used in a misleading context. The material must be acknowledged as Crown copyright and the title of the document/publication specified.

Contents	Page
Introduction	1
Background	2
Main findings	4
Recommendations	7
School based arrangements for standardisation and moderation	8
Key stage 2/3 cluster arrangements	10
Key stage 3 external moderation and verification	13
Support for assessment	16
An overall school perspective	17
Case studies	18

Introduction

- 1 This report is published in response to the annual Ministerial remit to Estyn for 2009-2010. It contains an evaluation of the arrangements to assure the consistency of teacher assessment in the core subjects in key stage 2 (KS2) and key stage 3 (KS3).
- 2 Since 2005, there have been no statutory tasks or tests in Wales at KS2 and KS3 and pupils' standards of attainment at the end of the key stages are based only on teacher assessments. The reasons behind the decision to discontinue statutory testing included:
 - the high reliability of teacher assessments;
 - the perception that the tasks and tests provided no significant additional information to that derived from teachers' own assessments of pupils' progress and attainment;
 - the time devoted to 'over preparing' pupils for the tasks and tests;
 - the potential to release time that could be directed to other more productive classroom activities;
 - the stress placed on pupils in undertaking the tasks and tests; and
 - the benefit of reducing the administrative burden of testing on teachers.
- 3 In order to undertake effective teacher assessment, schools were required in 2009 to have the following arrangements in place.
- 4 At KS2:
 - school based standardisation and moderation of examples of pupils' work;
 - KS 2/3 cluster group moderation of examples of pupils' work; and
 - end-of-key-stage teacher assessment for all eligible pupils.
- 5 At KS3:
 - school based standardisation and moderation of examples of pupils' work;
 - KS2 and KS3 cluster group moderation of examples of pupils' work;
 - the selection of sample evidence for external moderation of all subject departments; and
 - end-of-key-stage teacher assessment for all eligible pupils.
- 6 In addition, at KS3, procedures for teacher assessment were further supported by visits from external verifiers.

Background

- 7 The Welsh Assembly Government's (WAG) policy on assessment aims to ensure that there is confidence in the validity of end-of-key-stage teacher assessment outcomes at KS2 and KS3. These arrangements rely on having in place robust and reliable teacher assessment across schools throughout Wales. WAG's current assessment strategy has been developed following the decision to discontinue statutory end-of-key-stage testing for 11 and 14-year-old pupils from 2005 and to rely only on teacher assessments of pupils' attainment. The Department for Children, Education, Lifelong Learning and Skills (DCELLS) highlights that this decision 'places increased emphasis on teachers' understanding and application of the National Curriculum (NC) level descriptions'.
- 8 Following the abolition of tests at KS2 and KS3, schools have been required to engage in the process of school based and cluster based standardisation and moderation of pupils' work in order to support effective teacher assessments. These requirements are designed to help ensure that the best fit application of NC level descriptions to pupils' work is carried out accurately and consistently.
- 9 DCELLS's expectation for end-of-key-stage teacher assessment is that 'school based standardisation and moderation should form an integral part of each school's systems and procedures for ensuring high quality teacher assessment. This work should be part of the whole school year to ensure that each teacher is confident about national standards'.
- 10 In addition, although many clusters of schools have good systems to transfer assessment data as pupils move from the primary to the secondary school, many secondary departments do not use this information to plan timely intervention for pupils whose progress is too slow. Estyn's report on 'The impact of transition plans' published in June 2008 refers to the importance of transferring data, particularly in relation to pupils who have difficulties with basic skills or pupils who have had additional support in the primary schools through funded programmes to raise attainment and individual standards in education (RAISE).
- 11 Assessment is one of the weakest areas of work in schools. Estyn's evidence from school inspections across Wales consistently indicates that about a quarter of schools inspected each year have shortcomings in aspects of assessment. The Chief Inspector's Annual Report for the school year 2007-2008 judged that 'The lack of reliable and robust data on learner attainment is an issue that needs attention in a few sectors, particularly in the primary sector'. It further stated that 'At present, teacher assessments at key stage 1 (KS1) and KS2 are not consistent enough across Wales'.
- 12 The weakness of assessment in schools may be illustrated by evidence from individual school inspections over the last few years. In almost a third of primary school inspections in 2008-2009, the recommendations made by inspectors for school improvement refer to assessment issues. Amongst these, recommendations regarding the need for schools to improve the accuracy and consistency of assessment feature frequently.

- 13 In a few school inspection reports, references to assessment issues confirm that there is a lack of robust data on learner attainment. For example, the inspection report of one North Wales primary school states that 'Teacher assessments at the end-of-key-stage 1 and 2 are not subjected to robust external moderation. Comparison of standards of achievement in many pupils' books with end-of-key-stage teacher assessments does not reveal close correlation'.
- 14 Estyn's report on 'Best practice in mathematics for pupils aged 3 to 7 years' published in June 2009, states that 'Teacher assessment does not accurately reflect the proportion of pupils at age 7 that attain at level 3. Teachers sometimes assess work at level 2 when a level 3 would be more appropriate. This means that, in KS2, teachers do not challenge those pupils and this limits their progress'.
- 15 The inspection evidence quoted above indicates that assuring the accuracy and consistency of teacher assessment outcomes at the end of KS2 and KS3 is a significant challenge for schools. However, KS3 teachers have been supported in making accurate assessments of pupils' progress through the external moderation of core and non core subjects and verification of the general effectiveness of the assessment processes operating in each school.
- 16 The detail of this report includes an evaluation of the processes and outcomes of school based and cluster based standardisation and moderation in KS2 and KS3. This report also evaluates the overall effectiveness of the external processes of moderation and verification in KS3.
- 17 The report draws on:
 - information from Section 28 reports on the primary and secondary schools inspected during the period September 2006 – July 2008;
 - discussions with headteachers and teachers during visits made by Her Majesty's Inspectors to 23 schools including 11 primary schools and 12 secondary schools;
 - a detailed questionnaire response from 37 additional schools representing a cross section of primary and secondary schools in 20 out of the 22 local authority areas in Wales; and
 - a range of additional KS3 evidence including discussion with the deputy chief external verifier, ongoing discussions with DCELLS staff responsible for assessment and a sample of external moderators' and verifiers' reports received from secondary schools.

Main findings

- 18 Since the abolition of NC tests, the Welsh Assembly Government's policy on assessment has been mainly focused on arrangements to ensure robust and reliable teacher assessment at KS2 and KS3 across schools throughout Wales.
- 19 However, evidence from Estyn's inspections of schools consistently indicates that assessment is one of the weakest areas of work in schools. There are also weaknesses in the use that secondary teachers make of assessment information from primary schools.
- 20 Our survey indicates that overall, primary and secondary school teachers are becoming more confident about their understanding of the characteristics of pupils' work that demonstrate the NC level descriptions. Nevertheless, different perceptions as to what constitutes appropriate standards in the core subjects at different levels in end-of-key stage 2 and key stage 3 assessments continue to exist.

Arrangements for teacher assessment at the end of KS2 and KS3

- 21 In the most effective primary schools, school based standardisation and moderation meetings are being used to help to achieve better consistency in teacher assessment. However, the frequency of these meetings varies considerably and in four in every 10 primary schools surveyed, these types of meetings only take place once a year. While most primary schools engage in some standardisation activities so that teachers can better understand how to apply the 'levels', teachers do not actually come together to moderate pupils' work towards the end of the key stage. (The glossary at the end of this report explains the difference between 'standardisation' and 'moderation'.)
- 22 Staff in schools often do not make full use of the time that is made available to them for end-of-key-stage assessment activities. Two additional school closure days were available for schools to use for curriculum and assessment purposes last year but in over a half of primary schools surveyed, these days were not used for internal standardisation and moderation.
- 23 Arrangements for assuring the accuracy and consistency of end-of-key stage 2 teacher assessments are mainly based on a system of cluster-group standardisation and moderation meetings at which staff from a cluster of schools in a given locality come together. All clusters surveyed have engaged in these activities and almost nine out of 10 schools involved state that the cluster process is working well. In a few school clusters, teachers say that the meetings and their outcomes are not effective and that the levels teachers award to examples of pupils' work are not agreed by all partners within the cluster.
- 24 In three quarters of primary schools surveyed, the confidence of teachers to award teacher assessment levels in line with NC level descriptions has improved through the cluster based process. However, schools state that it is extremely difficult to ensure that the cluster's shared understanding of standards is actually applied in the assessments that teachers subsequently undertake in each of the cluster schools.

- 25 Statutory assessment of Welsh as a second language in KS2 will take place for the first time in the summer of 2010. Only about a half of the primary schools surveyed have made any arrangements to include this subject in school based standardisation and moderation meetings and most schools have no plans to include it in cluster based arrangements (explained in the glossary).
- 26 All secondary schools surveyed have arrangements in place for school based standardisation and moderation meetings in the core subjects to assure the accuracy of teacher assessment within and at the end of KS3. Teachers meet at least once a year in all these secondary schools to discuss and reach agreement on the assessment and levelling of pupils' work.
- 27 The accuracy of end-of-key stage 3 teacher assessments is further supported through a system of external moderation of teachers' understanding of subject standards and the verification of school based assessment systems and procedures. Generally, these processes have been welcomed by secondary school teachers.
- 28 The verification process involves the scrutiny of secondary school based assessment systems and procedures. Secondary schools surveyed state that this external scrutiny of assessment systems and procedures is helpful and, in most schools, teachers are now confident in their systems and procedures for assessment.

Concerns about the arrangements for teacher assessment

- 29 DCELLS has produced considerable guidance over the past few years to support schools in undertaking assessment. Overall, primary and secondary schools take good account of the guidance. Nevertheless, over a third of primary schools surveyed had not used all the publications and in these schools it is unlikely that the current guidance regarding the revised curriculum and associated assessment arrangements is being fully applied.
- 30 Teachers are concerned that current arrangements for teacher assessment are not working as well as they should because teachers at both KS2 and KS3 continue to have different interpretations of 'levels'. Many school clusters have identified further work that needs to be undertaken before they have sufficient confidence in the accuracy and consistency of the teacher assessment in NC core subjects.
 - Staff need to plan what the cluster of schools will do to ensure coverage of all core subjects within a realistic timescale.
 - All relevant staff at primary schools need to attend the standardisation and moderation sessions organised to confirm a shared agreement on teacher assessment judgements.
 - Better guidance and training is required on how clusters should apply the 'best fit' model for NC level descriptions.
 - Teachers want to review and evaluate cluster work to ensure that it also focuses on the process of moderation rather than just on standardisation. This will support the consistency and reliability of end-of-key-stage teacher assessment.

- Teachers want to increase the involvement of local authority advisory staff who can support cluster moderation work in terms of general advice on effective procedures and in specific NC core subjects.

31 This survey raises other related concerns:

- There is doubt about the ability of schools to continue with the current system of standardisation and moderation meetings. The implementation of 'rarely cover' arrangements from September 2009 may mean that teachers will find it more difficult to attend school cluster meetings and assessment training sessions.
- The application of a 'best fit' assessment methodology does not aid the process of assuring the consistency of assessment at the end of the key stages.
- Advice from DCELLS to avoid subdividing NC levels or using level descriptions for the purposes of ongoing assessment within the key stages is not seen as helpful. School staff say that, without referring to NC levels during, as well as at the end of the key stages, it is very difficult to track the progress of pupils effectively or to set accurate targets for further progress.

Reliability of assessment outcomes

- 32 Assessment outcomes in KS3 are more reliable than at KS2 because there are more robust systems in place for assuring the accuracy of teacher assessments at KS3.
- 33 In KS2, confidence in the system of assessment, and in particular in the consistency of its application across different school cluster groups, is not as reliable as it should be because there is not, as yet, an effective process in place to moderate these judgements externally.
- 34 In addition, as it currently operates, the process does not actually verify that the levels finally awarded to examples of pupils' work at the end of key stage 2 are accurate across the full range of individual schools, clusters of schools and local authorities. This means that a level 2, or any given level, will not mean the same in different parts of the country.
- 35 At KS1, there are no processes in place to verify or to assure the accuracy of end-of-key-stage teacher-assessment outcomes. As a result, inaccurate teacher assessment outcomes at the end-of-key-stages 1 and 2 make the reliability of value added and benchmark measures between the key stages uncertain. This is an important shortcoming because it brings into question the robustness of national data and benchmarks and the use of data for calculating the value added at the next stage.

Recommendations

The Welsh Assembly Government should put in place arrangements to:

- R1 provide further clear advice to schools regarding the use of the 'best-fit' model of assessment and the flexible use of level descriptions within the key stages;
- R2 sample evidence from teacher assessments across a range of levels, in each cluster of schools at least, to ensure that they are accurate and reflect national standards; and
- R3 encourage local authorities, as part of their role in implementing the school effectiveness framework, to embed processes to secure better accuracy and consistency in teacher assessment, particularly at KS1 and KS2.

Local authorities should:

- R4 take an active role in supporting the cluster moderation process directly where clusters of schools meet in their local authority areas; and
- R5 embed processes that will secure better accuracy and consistency of teacher assessments at KS1 and KS2, as part of their role in implementing the school effectiveness framework.

Schools should:

- R6 arrange cluster moderation sessions at least once a year to moderate pupils' work and confirm teachers' understanding of standards in the core subjects;
- R7 ensure that teacher representatives from each school attend all cluster meetings held for the purposes of standardisation and moderation, using as required the school closure days that are available for this purpose;
- R8 ensure that all teachers are aware of and use the relevant outcomes of their own cluster standardisation and moderation meetings;
- R9 include the assessment of Welsh second language in school and cluster standardisation and moderation meetings; and
- R10 ensure that DCELLS guidance on assessment is taken fully into account.

School based arrangements for standardisation and moderation

Key stage 2

- 36 All primary schools surveyed hold standardisation and moderation meetings (explained in the glossary). In the meetings teachers discuss the standards of pupils' work and the NC levels that they award to such work.
- 37 However, in our sample of schools the frequency with which teachers hold these meetings varies considerably. In four in every 10 primary schools, these types of meetings only take place once a year, while in just over a half of schools, the process is undertaken at least once a term. In a very few schools, standardisation and moderation meetings are held as regularly as once a month.
- 38 Teachers in primary schools hold standardisation and moderation meetings at various times including during staff meetings or after school meetings designed for training purposes. Just under a half of primary schools hold these meetings during after school training sessions. However, it is uncertain whether after school sessions enable teachers to be focused enough in the time available to carry out this demanding and detailed work effectively.
- 39 In the last year, two additional school closure days were available for schools to use for curriculum and assessment purposes. However, our survey indicates that in over a half of primary schools these days were not used for internal standardisation and moderation.
- 40 In nearly all primary schools surveyed, all teachers are involved in standardisation and moderation meetings where standards of pupils' work are discussed. In a majority of primary schools, headteachers or core subject co-ordinators take the lead in these meetings. Most of the time in these meetings is spent discussing pupils' work in English while the least time is spent on pupils' work in science.
- 41 In most primary schools, the outcome of standardisation meetings has been the production of school subject portfolios. These generally contain samples of assessed pupils' work from all KS2 year groups for all relevant levels. For each level, the selected pupils' work usually contains commentaries from teachers that identify the characteristics within the sample that show links to particular level descriptions. All primary schools that have produced portfolios of work concentrate on pupils' work in English or Welsh, with many schools also producing a portfolio of pupils' work in mathematics. Few schools have produced portfolios of pupils' work in science.
- 42 As a result of standardisation and moderation meetings in the primary schools surveyed, most teachers state that they are more confident about applying NC level descriptions accurately in English or Welsh. Many teachers state they are also confident to apply the level descriptions accurately in mathematics and science.
- 43 In a few schools, teachers have not yet started to put together portfolios of assessed pupils' work and to use them to inform and confirm their understanding of standards in relation to the NC level descriptions. Also, many KS2 teachers do not appreciate the difference between standardisation and moderation activities. As a result, most

primary schools engage in standardisation activities but not actually in the moderation of pupils' work to support best fit judgements towards the end of the key stage.

Welsh as a second language

- 44 Statutory assessment of Welsh as a second language in KS2 will take place for the first time in the summer of 2010 and teachers will assess pupils' attainment against the NC level descriptions.
- 45 Most primary schools that we surveyed are aware of this change but only about a half has made any arrangements to include the subject in school based standardisation and moderation meetings. About a quarter of the schools we surveyed are planning to include Welsh as a second language as a priority in their school development plans for next year but only a very few are planning to produce portfolios of pupils' work in relation to NC level descriptions.

Key stage 3

- 46 All secondary schools surveyed have detailed arrangements for school based standardisation and moderation meetings in the NC core subjects to assure the accuracy of teacher assessment within and at the end of KS3.
- 47 NC core subject teachers regularly meet to share examples of pupils' work and to discuss the characteristics of the NC levels that the work illustrates. A majority of this standardisation work is undertaken after school in designated time either during staff meetings or during 'twilight' in-service training sessions. In addition, in the last year, standardisation and moderation activities have been undertaken in three quarters of the schools during the school closure days available for this purpose. In all the secondary schools surveyed, teachers meet at least annually to discuss examples of pupils' work in relation to the NC level descriptions and, in many of these schools where the practice is most effective, teachers meet each term.
- 48 These core subject meetings are generally attended by all subject staff with senior staff in attendance in about a quarter of the schools. Subject Heads of Departments lead and coordinate the standardisation and moderation work in many schools and all schools surveyed have produced subject portfolios of assessed work in one or more of the NC core subjects.
- 49 These departmental subject standardisation portfolios contain samples of pupils' work along with teacher commentaries that demonstrate the characteristics of particular level descriptions. Many portfolios contain samples of pupils' work for Years 7, 8 and 9 and for all levels appropriate for KS3 pupils.
- 50 Secondary teachers state that the school based standardisation and moderation meetings have generally improved their confidence to assess pupils' work accurately in line with the NC level descriptions. All teachers in the mathematics departments surveyed are confident that their assessments of pupils' work are accurate. Nearly all teachers of English and most science teachers are also confident in assessing the standard of pupils' work accurately in relation to NC level descriptions.

Key stage 2/3 cluster arrangements

- 51 Arrangements to ensure a shared understanding of national standards between primary and secondary schools are based around a system of cluster group standardisation and moderation.
- 52 All clusters of schools surveyed engage in standardisation processes, but far fewer engage in the process of moderation to confirm end-of-key-stage best fit judgements on pupils' levels of attainment in the NC core subjects. Almost nine out of 10 schools surveyed state that the cluster process is working well with strong links being forged between primary and secondary schools. These strong links are facilitating open discussions between teachers that are helping to clarify judgements about pupils' work. However, a few schools state that cluster standardisation and moderation is in its infancy and they need far more time to develop effective arrangements that will produce secure agreement on end-of-key-stage outcomes and confirm alignment to national standards.
- 53 DCELLS's expectation of schools in respect of cluster arrangements for assessment states that 'planning for cluster groups to undertake moderation of learners' work should be across the school year. It is for individual KS2 and KS3 cluster groups to decide the timing and frequency of their meetings. Cluster group moderation should only need to take place once during the school year'.
- 54 We found that the frequency of standardisation and moderation meetings varies considerably between clusters of schools. Just over a half of schools surveyed hold cluster sessions once a year while just under a third attend a meeting each term. The majority of schools hold these cluster sessions on the school closure days available for the purpose while a quarter undertake standardisation and moderation work during 'twilight' in-service training sessions. Not all school staff attend cluster meetings. In primary schools, teachers who attend are usually Year 6 teachers or core subject co-ordinators while in secondary schools teachers with core subject responsibilities at KS3 and heads of department usually attend.
- 55 A quarter of school clusters have yet to produce agreed portfolios of examples of pupils' work as a result of cluster meetings. In the clusters that have produced portfolios, the majority have a portfolio of agreed examples of pupils' work in English with fewest having portfolios of pupils' work in science. Less than a third of the portfolios that have been produced contain samples of work from all KS2 year groups and about a half only contains work from Year 6 pupils.
- 56 Overall, the cluster based standardisation and moderation process is having a positive effect in schools. About three-quarters of schools surveyed state they are more confident about the accuracy and consistency of teacher assessments in the core subjects as a result of the process.
- 57 However, many schools have identified further work that they need to undertake before they will have sufficient confidence in the accuracy and consistency of the teacher assessment levels awarded in all NC core subjects. A few schools state that the structure of the cluster standardisation and moderation meetings needs to be improved and, in these schools, they are not clear what the outcomes of the meetings and discussions about pupils' work should be.

58 Currently, there are a number of issues that the cluster based system we surveyed has yet to address in enough detail:

- Not all school clusters have clear plans of what they aim to do and how the outcomes of their discussions about the standards of pupils' work should be used. Just over one in 10 clusters state that their current arrangements are not working well.
- Most school clusters have no current plans to include Welsh as a second language in cluster based arrangements for standardisation and moderation. This fails to take advantage of teachers' expertise within the clusters. It will also make it difficult to assure the accuracy of teacher assessment outcomes within each school when the statutory assessment of Welsh as a second language in KS2 takes place in the summer of 2010.
- Many clusters of schools contain a large number of primary schools and cluster co-ordinators have found that arranging meetings for all of these schools has often been difficult. In a very few cases, not all primary schools have attended the cluster standardisation and moderation meetings held. This means that in these cluster groups they cannot be certain that all the schools have the same shared understanding of the NC level descriptions and their application to pupils' work.
- Even when all cluster schools have been present for the standardisation and moderation meetings, ensuring that the cluster's shared understanding of standards is applied within each of the cluster's schools is extremely difficult to achieve and there is no subsequent check on this.
- Schools and clusters do not differentiate enough between activities that involve standardisation and activities that involve the moderation of the actual work that is being assessed at the end of the key stage. A greater emphasis on moderation would help to offset the distrust that exists amongst secondary teachers in a few school clusters regarding the levels awarded to pupils' work at the end of KS2. This distrust is related to the perception that pupils' work has not been moderated on the basis of a shared understanding of standards.
- Not enough clusters of schools think about when it would be most useful to meet for moderation. The most effective timing for meetings would be when primary schools are confirming KS2/KS3 teacher assessment judgements about pupils' attainment. This would make the moderation meetings more relevant and help primary school teachers provide accurate and agreed judgements about pupils' attainment. This timing would also give all schools within the cluster confidence in the levels awarded.
- Three-quarters of primary schools surveyed state that the current cluster based moderation process has increased their confidence to award teacher assessment levels in line with NC level descriptions. However, towards the end of KS2 in most school clusters, the moderation of pupils' work does not reflect enough examples of Year 6 pupils' work. No check is made on the actual end of KS2 teacher assessments across the range of primary schools in each cluster or

even of a sample of teacher assessments from each cluster. Therefore, teacher assessment results at the end of KS2 are not actually verified or sampled in any way. In addition, at KS1, there are currently no processes in place to assure the accuracy of end-of-key-stage teacher-assessment outcomes.

- In a few school clusters, the way that schools interpret the 'best fit' model of applying level descriptions when assessing pupils' work is incorrect.
- End-of-KS1 and 2 outcomes, value-added and school benchmark measures are used in a number of ways to make judgements about national standards and to compare the standards that individual pupils, cohorts, schools and local authorities achieve. Inaccurate teacher assessment outcomes at the end of KS1 and KS2 make the reliability of value added measures between the key stages uncertain. This is an important shortcoming because it also brings into question the robustness of national data and benchmarks.
- In many clusters of schools, local authority (LA) advisory staff attend standardisation and moderation meetings to offer particular subject advice and support. The attendance of LA advisory staff often provides a steer to the meetings and can often emphasise the importance the LA attaches to these meetings.
- In only a very few LAs are advisory staff involved in attending standardisation and moderation events in all clusters in their areas. In these LAs, the attendance of advisory staff at meetings has a very positive impact on the consistency of application of NC level descriptions within and across clusters of schools (see Case Study 1). However, even in the best LAs where they provide outstanding support for the assessment process, there are no systems for inter local authority standardisation or moderation to take place.
- Currently, LAs have no statutory duties in relation to the assessment processes in schools. LA advisory staff are therefore largely excluded from dealing with assessment issues in primary schools although, historically, many have carried out a very rigorous check on the accuracy of teacher assessment levels in their schools.

Key stage 3 external moderation and verification

- 59 Arrangements for securing the accuracy and consistency of end of KS3 teacher assessments are additionally strengthened by a system of external moderation of subject departments' understanding of national standards and the verification of school based assessment systems and procedures. Subject teams of external moderators review and confirm school departments' understanding of NC level descriptions based on teachers' commentaries of selected examples of pupils' work. External verifiers also evaluate school assessment systems and procedures.
- 60 The arrangements for secondary schools are now well advanced. All core subject departments (and just over 50% of non core subject departments) in mainstream schools have been moderated. At the time of writing, the majority of schools have also had their assessment systems and procedures verified.
- 61 The evidence now available from external moderation and verification has confirmed a mixed picture in terms of quality, ranging from schools where teacher assessment systems are outstanding to schools or individual subject departments which do not demonstrate high quality teacher assessment. Teacher assessment in science is much weaker than teacher assessment in the other core subjects.

Moderation

- 62 Our school sample indicates that the process of external moderation has been generally welcomed by secondary schools and has generally increased the confidence of KS3 teachers to assess accurately against the NC level descriptions. The process has improved the accuracy and consistency of teacher assessments between subject departments in schools within the same LA and across different local authorities.
- 63 All secondary schools surveyed were generally clear about the evidence they needed to provide for external moderation purposes in core subject departments. However, about a quarter of science departments state that there were a few initial issues with guidance about the evidence they needed to provide. Also, a few subject departments state that the requests for materials for external moderation were considered unreasonable and that the process had increased their workload.
- 64 Following an evaluation of school departments' evidence, external moderators provide reports about the accuracy of teachers' understanding of national standards. Many core subject departments found these reports very useful. However, a few science departments stated that the reports provided only limited reasons for the external moderators' non agreement of levels. Also, a few departments found that moderators' reports were too general in nature to enable teachers to make improvements.

- 65 In most secondary schools surveyed, moderators required subject departments to provide further evidence for a number of reasons. These included:
- too little evidence provided by teachers for moderators accurately to judge their understanding and application of standards;
 - a lack of teachers' annotation on pupils' work to indicate why it was judged to be characteristic of a particular level;
 - use of inappropriate subject content outside the KS3 programmes of study; and
 - indication from the initial evidence presented that teachers do not have a sound understanding of some/all level descriptions.
- 66 Nearly all English departments in the secondary schools surveyed stated they are now more confident that their teacher assessments are accurate. Also the external moderation processes have prompted a number to make standardisation a more regular practice in their schools. Most mathematics departments and many science departments also stated that they are better at assessing pupils' work against the NC level descriptions.
- 67 In just over a third of the secondary schools surveyed, they have changed assessment practice in one or more core subject departments as a result of moderators' feedback. Some of the changes made include the following:
- teachers place less reliance on past subject test papers;
 - subject departments discuss pupils' work more regularly and identify the characteristics of the level descriptions that the work illustrates;
 - in English, teachers use better oral and reading assessments;
 - in mathematics, teachers put more emphasis on collecting and judging evidence of pupils' ability to apply their mathematical skills in other subjects across the curriculum and to real life problems; and
 - in science, teachers use evidence from more extended tasks in order to illustrate the broader scientific knowledge pupils have.
- 68 Where core subject department teachers and heads of department have drawn up action plans following external moderator's feedback they are confident that the issues identified for future improvement will be dealt with effectively.

Verification

- 69 The verification process involves an external scrutiny of secondary school assessment systems and procedures via a school visit.
- 70 Generally, the process of external verification, as with moderation, has been welcomed by secondary schools and teachers state that they feel reassured that it

has improved the accuracy and consistency of teacher assessments between subject departments, schools and local authorities. The general view is that an external scrutiny of assessment systems and procedures is a helpful process.

- 71 Most secondary schools surveyed state that the evidence they were required to provide for verification purposes was clear and reasonable. Many secondary school teachers state that verifiers discussed local school issues with them and gave good guidance on sources of best practice. Verifiers' feedback reports to subject departments were felt by most to be clear and helpful. However, in the early stages of verification, a very few schools were still waiting for reports several months after the verifier's visit. This was not helpful for the schools concerned.
- 72 As a result of the verification process, most secondary schools state they are confident that their systems and procedures for assessment are working well. However, just over a third of schools have changed their assessment practices as a result of verifiers' feedback. Some of the more common changes made include:
- revising assessment policies to better reflect school practice;
 - developing learner profiles;
 - adjusting existing learner profiles to exemplify subject attainment targets better; and
 - revising standardisation and moderation processes.

Support for assessment

- 73 DCELLS has produced considerable guidance over the past few years to support schools on how they assess, record and report on pupils' progress. The materials available to schools during the survey included the following publications:
- Ensuring consistency in teacher assessment: Guidance for key stages 2 and 3;
 - Making the most of learning;
 - Optional assessment materials;
 - Exemplification materials linked to the revised school curriculum (Curriculum 2008);
 - Statutory assessment arrangements for the school year 2008-2009 – key stages 2 and 3; and
 - Question and answer guide to key stage 2 and 3 assessment arrangements, school year 2008-2009
- 74 Overall, primary and secondary schools take good account of the guidance and many primary and all secondary schools have used the publication 'Ensuring consistency in teacher assessment: Guidance for key stages 2 and 3' to support their standardisation and moderation arrangements.
- 75 However, over a third of primary schools surveyed had not used the publication 'Making the most of learning' or the exemplification materials linked to the revised curriculum Orders. These are key documents as they provide advice on changes to the content and presentation of the revised curriculum Orders and on important assessment issues. As a result of not using these documents, a third of primary schools may not be applying the current guidance regarding the revised curriculum and associated assessment arrangements.
- 76 More secondary schools than primary schools use the above list of publications. A high percentage of secondary schools use all the publications, with the exception of the publication 'Question and answer guide to key stage 3 assessment arrangements, school year 2008-2009'.

An overall school perspective

77 We asked the schools surveyed to give us an overall perspective of Welsh Assembly Government's assessment strategies. The following provides a summary of the positive responses we received:

- The current strategies undertaken in schools to secure accurate end-of-key-stage teacher assessment have improved links between KS2 and KS3 teachers. About three quarters of the schools surveyed state that there is now a better shared understanding of assessment issues between teachers in the different key stages. In addition, teachers in those schools state that they have improved confidence in assessing the standards of pupils' work and in making 'best fit' end-of-key-stage judgements of overall subject levels.
- In secondary schools, the external processes of moderation and verification have provided valuable perspectives of schools' assessment systems and procedures.

78 However, school staff state that they have areas of concern about the current assessment arrangements:

- There is doubt from about a half of both primary and secondary schools surveyed that they will be able to continue with the current system of standardisation and moderation meetings. These processes are very time consuming and unless the additional INSET days currently available to schools continue, it will be difficult for teachers to find the time to attend relevant discussions and meetings. In a few schools, they are planning to schedule standardisation and moderation sessions into their annual departmental and school plans and timetables. However, the implementation of 'rarely cover' arrangements from September 2009 may mean that teachers will find it more difficult to attend cluster meetings in the future and assessment training sessions.
- The application of a 'best fit' assessment methodology is not specific enough to ensure an accurate and consistent approach to the awarding of levels at the end of the key stages. In secondary schools in particular, this has contributed to a lack of confidence particularly amongst non specialist core subject teachers.
- DCELLS's advice to avoid subdividing NC levels or using level descriptions within the key stages does not aid the assessment process. Schools state that, without referring to pupils' work in relation to NC levels within the key stages, it is difficult to track the progress of pupils effectively or to set accurate future targets. Many schools have subdivided levels in an agreed way in order to better identify the progress pupils make. This can sometimes work well, particularly when schools compare the levels pupils have attained at the end of the previous key stage with those that they are currently attaining. However, the subdivisions of levels made in schools are arbitrary and cannot be used between schools to make any sort of comparisons of pupils' progress.
- There is still a concern in many school clusters that teachers interpret NC levels differently in KS2 and KS3, with consequently different expectations of the work that pupils need to do to attain a particular NC level.

Case studies

The case studies below illustrate good practice in standardisation and moderation activities to secure accurate end-of-key-stage teacher assessment

Case study 1: How Newport local authority (LA) has developed an effective standardisation and moderation process in its clusters of schools

Context: In Newport, the cluster standardisation and moderation process is well supported by local authority advisory staff. Core subject specialist advisory staff attend meetings in all school clusters held for the purposes of standardisation and moderation. Separate meetings are held in each cluster for each NC core subject and are attended by teachers from each primary school and at least one teacher from the relevant core subject department in the secondary school. The primary teachers who attend usually teach in Year 6 classes or are the subject co-ordinators for the relevant core subject being discussed.

Strategy: Moderation sessions are held in each cluster of schools in the LA with a focus on the NC core subjects of English, mathematics and science. The sessions are held over a series of half days within a two week period.

LA advisory staff are present at all sessions and they use these opportunities to disseminate good practice in the specific NC core subject areas. In each session, LA advisory staff refer appropriately to recent DCELLS national curriculum guidance documents for KS2 and KS3.

Each moderation session begins with advisory staff referring to exemplar pupils' work included in DCELLS core subject guidance documents. These samples show clearly how level descriptions can be used when making judgements about which NC level best describes a learners' overall performance at the end of KS2 and KS3.

Following this, teachers use their own pupils' work to promote discussion on NC levels and what exemplifies achievement at a range of levels.

Teachers from the primary schools use pupils' work at levels 3, 4 and 5. Teachers from the secondary school use pupils' work at levels 4, 5 and 6.

Action: As the moderation sessions proceed, teachers reach general agreement on pupils' work that illustrates achievement at specific levels. Pupils' work that illustrates achievement on the threshold between levels is more difficult to agree.

Outcomes: All teachers attending the sessions find them very useful. Many teachers made positive comments involving the:

- valuable lead taken by local authority advisory staff in disseminating good practice in the NC core subjects and for providing clear guidance on the levels awarded to pupils' work;
- productive discussions that take place between teachers from different schools and phases about pupils' work;

- general agreement about the types of pupils' work that exemplify characteristics from the different NC levels;
- confidence that they gain from the process in understanding the level descriptions and applying this understanding to the assessment of pupils' work;
- clarification of pupils' attainment at the thresholds between the different NC levels; and
- sharing of assessment expertise.

This type of well organised and well supported cluster moderation is:

- enabling schools to make accurate and consistent teacher assessment judgements about pupils' work, particularly at the end of KS2;
- providing consistency in the assignment of levels to pupils' work across clusters through the involvement of LA core subject advisory staff; and
- ensuring the alignment of all schools within the LA to the application of national standards.

Case study 2: How Brecon High School ensures that teachers' judgements about pupils' attainment at the end-of-key-stage 3 are accurate

Context: In Brecon High School, teachers in the English department have agreed a process to ensure that they take a consistent approach to assessment.

Strategy: A process of ongoing standardisation and moderation is undertaken by all teachers in the English department.

Action: Each half term, two or three pieces of unmarked work from Year 9 pupils are distributed to all teachers in the English department to be marked and the characteristics of NC level descriptions that the work illustrates are identified. During planned timetabled standardisation and moderation sessions, teachers discuss and identify the particular strengths exemplified in each pupils' work.

Towards the end of the key stage, a common assignment is undertaken by all pupils in a year group. These assignments, when marked, are checked by the Head of Department to ensure that marking is undertaken in a consistent manner.

At the end of Year 9, pupils' work is collected into a learner profile with each scrutinised and cross moderated with other teachers in the English department before final NC 'best fit' levels are decided.

Outcomes: All teachers of English in the school are confident in the assessment process. They are provided with clear guidelines and assessment criteria for the pupils' tasks that they assess.

The work of all Year 9 pupils is discussed and moderated effectively by all teachers of English in planned sessions. The cross moderation of pupils' work at the end of Year 9 supports accurate and consistent teacher assessment which represents the combined judgements of all teachers in the English department.

Case study 3: How Monmouth Comprehensive School ensures that best practice in standardisation and moderation is disseminated between the different subject departments

Context: Monmouth Comprehensive School has produced a best practice guide to ensure that all subject departments are aware of up-to-date advice from DCELLS regarding standardisation and moderation and the outcomes of external moderation and verification at the school.

Strategy: By fully disseminating all aspects of the standardisation, moderation and verification processes in the different subject departments, the school is ensuring a consistent approach to end of KS3 teacher assessment.

Action: The production of a best practice guide provides all teachers in the school with detailed information about the KS3 teacher assessment processes in place. The guide contains sections on:

- the planning of learning and assessment;
- consistency in assessment;
- internal standardisation and moderation processes;
- individual subject department assessment processes;
- up-to-date news on assessment from DCELLS;
- a schedule for external moderation and verification at the school;
- external moderation subject outcomes;
- external verification outcomes;
- further school assessment developments as a result of the external processes of moderation and verification; and
- planned professional development activities for teachers that will support assessment.

This guide is further supported by school policies relating to assessment. These policies include:

- Assessment for learning;

- Marking and homework;
- Supporting the learning of all students;
- Key skills; and
- Standards and consistency

Outcomes: The best practice guide and related policy documents provide all teachers with detailed information regarding the assessment processes in the school. This information gives teachers confidence in the school systems and a clear knowledge of what is happening in the separate subject departments. The sharing of best practice between departments is having a positive effect in the school and best practice 'hot spots' are regular discussion items in departmental meetings and training sessions. This results in improved consistency in assessment within and between subject departments in the school.

Glossary

achievement	How well pupils are doing in relation to their ability and by the progress they make.
attainment	How well pupils are doing as measured in national tests and in the qualifications or credits they gain.
benchmark data	This refers to the assessment information that schools use to compare their performance with that of other schools.
clusters of schools	Local authorities use different terms to describe groups of schools that comprise the secondary school and its feeder primary schools. For the sake of consistency, the term 'cluster' is used in this report.
moderation	Moderation occurs at the end of a key stage where a 'best fit' judgement on an individual learner's level of attainment is made.
national curriculum core subjects	These are English, Welsh, mathematics and science
standardisation	This involves a process of using samples of the work of the same learner or of different learners to enable teachers to reach agreement on levels of attainment by confirming a shared understanding of the characteristics of a level.
value added	This is a measurement of the amount of improvement that a school has brought about in a pupil over time. It is the relative advantage that a school gives a pupil, after taking into account the pupil's ability.

References

- Making the most of learning – DCELLS 2008
- Ensuring consistency in teacher assessment – DCELLS 2008
- Statutory assessment arrangements for the school year 2008-2009 (Primary) – DCELLS 2008
- Statutory assessment arrangements for the school year 2008-2009 (key stage 3) – DCELLS 2008
- Question and answer guide to key stage 2 teacher assessment arrangements, school year 2008-2009 – DCELLS 2008
- Question and answer guide to key stage 3 teacher assessment arrangements, school year 2008-2009 – DCELLS 2008

Explanation of words and phrases used to describe the report evaluations

The words and phrases used in the left hand column below are those that we use to describe our evaluations. The phrases in the right hand column are the more precise explanations.

nearly all	with very few exceptions
most	90% or more
many	70% or more
a majority	over 60%
half/around half	close to 50%
a minority	below 40%
few	below 20%
very few	less than 10%

The remit author and survey team

Mike Maguire, HMI	Author
Ann Jones, HMI	Survey team member
Peter Roach, AI	Survey team member
Alwyn Thomas, AI	Survey team member